CESifo

WORKING Papers

The Efficiency of Educational Production: A Comparison of the Nordic Countries with other OECD Countries

Peter Bogetoft
Eskil Heinesen
Torben Tranæs

CESifo
Center for Economic Studies & Ifo Institute

The Efficiency of Educational Production:
A Comparison of the Nordic Countries
with other OECD Countries

Abstract

Focusing in particular on upper secondary education, this paper examines whether the relatively high level of expenditure on education in the Nordic countries is matched by high output from the educational sector, both in terms of student enrolment and indicators of output quality in the form of graduation/completion rates and expected earnings after completed education. We use Data Envelopment Analysis (DEA) to compare (benchmark) the Nordic countries with a relevant group of rich OECD countries and calculate input efficiency scores for each country. We estimate a wide range of specifications in order to analyse different aspects of efficiency. In purely quantitative models (where inputs and outputs are expenditure and number of students at different levels of the educational system) and in models where graduation or completion rates are included as an indicator of output quality, Finland is the most efficient Nordic country (often fully efficient), whereas Sweden and especially Norway and Denmark are clearly inefficient. However, using PISA test scores as indicators of student input quality in upper secondary education reduces the inefficiencies of these three countries. Also, when expected earnings after completed education are used as an indicator of output quality, all Nordic countries are estimated to be fully efficient (or nearly so).

Peter Bogetoft
Copenhagen Business School
Department of Economics
Porcelaenshaven 16 A
Denmark – 2000 Frederiksberg
pb.eco@cbs.dk

Eskil Heinesen*
Rockwool Foundation Research Unit
Sølvgade 10
Denmark – 1307 Copenhagen K
esh@rff.dk

Torben Tranæs
SFI – Danish National Centre for Social Research
Herful Trolles Gade 11
Denmark – 1052 Copenhagen K
ttr@sfi.dk

*corresponding author

1. Introduction

Efficiency in education has always been important, but new dimensions will be added to its importance in the future due to increasing globalisation. Firms will increasingly be able to cover their requirements for a trained workforce by employing immigrants educated in other countries, and young people will increasingly be able to acquire their education abroad. If a given educational programme can be taught more efficiently abroad, arguments for having the taxpayers finance domestic supply of that educational programme will be weaker.

There is still no truly international market for education, but the trend is clearly moving in that direction, even for countries such as the Nordic where the native languages have a relatively small number of speakers. A somewhat recent feature in the Nordic countries is that globalisation has spread to the area of vocational training. First and foremost, this is a result of the closer integration of EU labour markets and the accession of nations in Eastern and Central Europe to the Union. These developments are taking place concurrently with the long-standing efforts of the EU to harmonise educational programmes in the various member states, thus making it easier for people to take courses outside their home countries. This makes completely new demands of each country's educational system in general, including upper secondary education.

Using Data Envelopment Analysis (DEA) based on OECD and Eurostat data for 2010, this paper focuses on the efficiency of the Nordic countries' educational sectors by means of comparison with the most relevant (the wealthier) OECD countries, and more narrowly the countries of Northern Europe. There are at least two reasons why it is interesting to focus on the Nordic welfare states. First, they are characterized by very high levels of public spending on education (accounting for 5.6-6.7% of GDP, much more than the OECD average of 4.5%), and also high levels of total educational spending including private spending:. Whereas average expenditure on primary, secondary and tertiary education in OECD countries is 5.9% of GDP, Denmark and Norway spend a total of 6.7 and 6.8% of GDP on education, and are among the countries that spend most, the others being Canada, the USA and New Zealand; Finland, Iceland and Sweden spend 6.1, 6.2 and 5.7%, respectively (OECD, 2013, Table B2.3). Since GDP per capita is high in the Nordic countries educational expenditure per capita is also very high. Second, the Nordic welfare states are characterized by high levels of spending on welfare programmes which reduce social problems and benefit public health (also for children and young people) and therefore promote an environment in which one may expect high efficiency in the education sector and high returns to human capital investment. Our analyses focus primarily on educational efficiency at the upper secondary school

level, but also take into account educational production more generally. We focus on the upper secondary level of education since most previous studies on the efficiency of educational production have focused on the primary and lower secondary level (e.g., Hanushek and Kimko, 2000; Gupta et al., 2002; Hanushek and Luque, 2003; Giménez et al., 2007; Miningou and Vierstraete, 2013; Coco and Lagravinese, 2014), or on the higher level (e.g., Barros et al., 2011; Lu, 2012; Zoghbi et al., 2013). Herrera and Pang (2005) analyse efficiency of overall educational spending for a large sample of developing countries.

An international comparison of educational efficiency is no simple matter, and the results will inevitably involve a significant degree of uncertainty even if we define the group of comparison countries narrowly. Besides the general problem that principles of assessing inputs and outputs may differ between countries, analysis of educational efficiency is difficult due to the fact that the educational systems differ a lot between countries, both regarding primary, secondary and tertiary education. For instance, the English-speaking countries and the countries bordering the Mediterranean have an upper secondary school system similar to the American system with one broad, unified and universal high school system, whereas the Nordic countries, Germany, Austria and Switzerland have a split at the upper secondary level into academic course programmes, which are geared towards preparation for further studies, and vocational training programmes, which lead directly to qualifications for skilled trades. Our approach, therefore, is to limit the comparison group as discussed above, and to present a number of estimates which, taken together, provide an impression of the relative efficiency of upper secondary programmes and the educational system more generally in the Nordic countries. We analyse models comparing the total number of students with total expenditure, but also models using quality indicators of educational output, including graduation rates, completion rates and expected earnings after completed education.

Not many DEA analyses exist that focus on international comparisons of upper secondary education. Based on data from OECD (2006), Verhoeven et al. (2007) conducted a DEA analysis of the relationship between expenditure per student in lower and upper secondary education on the one hand and upper secondary graduation rates on the other. Their results indicate that potential savings of Denmark, Finland and Sweden are considerable. Afonso and St. Aubyn (2005, 2006) carried out DEA analyses focussing on the last years of lower secondary school where the output is PISA scores. Afonso and St. Aubyn (2006) apply advanced two-stage methods suggested in Simar and Wilson (2007) in an analysis for 25 countries. The first stage is a DEA analysis where the output is average PISA scores and where two inputs measured in physical units are used: the number of

teachers per student in secondary school and the number of school instruction hours per year for 12-14-year-olds. In a second stage the estimated efficiency scores from the first stage are corrected for two important non-discretionary variables: GDP per capita and the share of the 35-44-year-olds in the population having at least an upper secondary education. The correction is based on regressions of the first-stage efficiency score estimates on these two variables, which are both clearly significant with positive effects on efficiency, and which are interpreted as indicators of family economic background and the educational level of the parents, respectively. This correction is very important for efficiency ranking. For instance, the Nordic countries with high GDP per capita and high levels of education are much less efficient after the correction: Finland and Sweden, which are fully efficient before the correction, are ranked 8 and 15 among the 25 countries after the correction, and Denmark and Norway are moved from positions near the middle of the distribution to the very bottom. The contribution of Afonso and St. Aubyn is important for several reasons. First, it is important to investigate the importance of indicators of parental background in country-level efficiency comparisons, since the large literature using individual-level data to investigate determinants of student test scores typically find large and significant effects of parental income and especially of parental education. For an example of this in a context of an efficiency study, see Bogetoft and Wittrup (2011). Second, test scores such as PISA and TIMSS scores are found to be important indicators of educational quality of countries. For instance, Hanushek and Kimko (2000) find that country differences in TIMMS scores are significantly correlated with differences in economic growth.

Sutherland et al. (2007) also use PISA scores as outputs, both average scores and a measure of homogeneity. As inputs they use in their preferred specification cumulative spending per student and an index of students' socioeconomic status. Using DEA methods on a sample of OECD countries, they find that Finland is among the most efficient countries, whereas savings potentials are substantial in Sweden and Denmark, and especially in Norway.

The present paper focuses on efficiency of upper secondary education. Therefore, we are not able to use PISA scores as an indicator of output quality, since they measure skills of 15-year-olds, i.e. at the end of lower secondary schools. However, in one of our analyses for upper secondary education, we use PISA scores as an input quality indicator, since higher PISA scores indicate higher quality of the student input which upper secondary schools receive from lower secondary schools. The models we consider in this paper are more in line with the approach in Verhoeven et al. (2007) than with Afonso and St. Aubyn (2006) in the sense that we use graduation rates as

output quality indicators in some of our models, and our dominant input variable is total expenditure (per student). There are pros and cons of using total expenditure instead of physical inputs. One advantage is that by using expenditure we compare countries not only according to their technical ability to transform given physical resources into educational output, but also according to their ability to find the cheapest mix of production factors. Another advantage is that it is important to reduce the number of inputs (and outputs) when the number of countries in the analysis is small.

The main contribution of this paper is that we use different DEA models to shed light on efficiency in especially upper secondary education in the Nordic countries relative to other wealthy OECD countries, using different specifications of inputs and outputs, including specifications where the output is expected earnings after completed education (corrected for expected employment rates). Earnings is the most important outcome in the vast literature on individual returns to education, but no studies has (according to our knowledge) used this as an outcome in DEA models of educational efficiency. Advantages of using this outcome are that it is more robust to institutional differences between countries than, e.g., graduation rates and that it is a measure of productivity, which is the most important purpose of education from an economic perspective. There are also disadvantages, especially that the current levels of earnings for different education groups depend on educational investments over many years for many cohorts and also on institutional features of the individual countries other than those narrowly related to the education system. However, we argue that earnings are an important indicator of educational outcome quality which can be used as one of a range of indicators.

International comparisons of efficiency in education production are very challenging because of the limited data available and the differences in educations and education systems across countries. On the other hand, it is also very relevant, not the least in light of the growing internationalisation of education. We therefor think it is worthwhile to analyses the prevailing data, and to try to draw conclusions as to what can and cannot be learnt from the available data.

The limited data also means that strong econometric approaches cannot be introduced. Instead we have relied on the DEA approach that allows us insight into the relative position of the countries when we involve multiple inputs and outputs in the evaluations. Indeed, our use of DEA in this paper serves primarily as an analytical tool to facilitate comparisons and to derive insight about the potential strengths and weaknesses of individual countries. The use of DEA is not intended to derive authoritative scores for the individual countries on which policy can directly be informed.

2. Data on educational expenditure and outputs

Most of the data used in our analysis come from the OECD's database, which among other things provides the basis for the annual publication *Education at a Glance*. The data used for most of the variables are the same as those used as the basis for *Education at a Glance 2012*, but some information which was not available there has been drawn from the 2013 issue. In addition, data concerning rates of employment and earnings related to levels of education have been taken from Eurostat (2014).

The first four columns in Table 1 show expenditure per student in primary school, in lower secondary school, in upper secondary school, and in tertiary education. Expenditure is in US dollars and are Purchasing Power Parity (PPP) corrected. The table shows that Denmark, Norway and Iceland (together with other countries including the USA) spend relatively heavily per student in primary school. At the upper secondary level Norway spends considerably more than the other Nordic countries. Expenditure per student in tertiary education is high in Sweden, Norway and Denmark, but USA is far above the other countries. Finland and Iceland tend to have rather low expenditure per student.

The next three columns in Table 1 indicate how large a proportion of a birth cohort is expected to graduate from upper secondary school, short-cycle courses of tertiary education and medium-cycle or long-cycle courses of tertiary education. The calculation of these graduation rates involves a considerable amount of uncertainty, and the fact that education systems differ to such a

Table 1. Expenditure per student, graduation rates, and earnings by levels of education (2010)

	Expenditure per student (in PPP-corrected USD)				Graduation rates (percent)			Earnings by level of education (in PPP-corrected EUR)		
	Primary	Lower secondary	Upper secondary	Tertiary	Upper secondary	Short tertiary	Medium and long tertiary	Lower secondary	Upper secondary	Tertiary
Denmark	11,166	11,078	10,996	18,556	86	9	50	31,949	36,975	46,287
Finland	7,368	11,338	7,739	16,569	93	0	49	29,429	29,102	38,830
Norway	11,833	12,505	14,983	19,269	87	0	42	29,315	35,515	50,309
Sweden	9,382	9,642	10,375	19,961	75	6	37	26,934	28,893	35,886
Iceland	10,099	9,778	7,934	9,939	88	2	60	19,673	29,296	33,919
Germany	6,619	8,130	11,287	15,711	87	14	30	21,447	35,441	59,726
Netherlands	7,917	11,708	11,880	17,849		0	42	26,953	34,089	51,310
Belgium	8,341			15,443				30,007	33,785	51,018
Luxembourg	16,494	19,202	19,443		70			27,826	37,746	65,685
UK	9,088	10,124	9,929	16,338	92	12	51	27,323	30,167	42,262
Ireland	8,219	11,069	12,731	16,420	94	22	47	29,229	32,826	45,100
Austria	10,080	12,442	12,737	14,257		12	30	22,153	34,411	54,434
Switzerland	10,597	14,068	17,013	21,577		16	31	29,908	39,874	64,863
France	6,373	9,111	12,809	14,642				22,609	25,556	38,203
Italy	8,669	9,165	9,076	9,562	83	1	32	23,645	31,099	41,892
Spain	7,446	9,484	11,265	13,614	80	16	30	22,334	27,020	37,415
Australia	8,328	10,273	9,916	16,074		16	50			
New Zealand	6,812	7,304	8,670	10,619		26	47			
Canada	8,262		10,340	20,932	81	29	36			
USA	11,109	12,247	12,873	29,201	77	11	38			

Sources: OECD (2012); Eurostat (2014); own calculations.

great degree also means that the classification of programmes into these two types of tertiary education may be debatable. For upper secondary education, Finland has the highest graduation rate (93%) among the Nordic countries, Sweden the lowest (75%), whereas it is 86 and 87% for

Denmark and Norway; for tertiary education Denmark and Iceland have the highest graduation rates among the Nordic countries.

The final three columns show average earnings for given levels of education, converted to euro using PPP correction. Compared with other countries, earnings for people with no education beyond the lower secondary level are very high in Denmark, and to some extent also in Finland and Norway. Earnings of people who have completed upper secondary schooling (but no tertiary education) are also rather high in Denmark and Norway compared to other countries. Earnings for people with tertiary education are higher in Norway than in the other Nordic countries, but they are even higher in many other European countries, e.g. Germany, Austria and Switzerland. Earnings dispersion, measured by the difference in average earnings between people with tertiary and lower secondary education, is very small in Finland and Sweden (about 9,000 euro), it is about 14,000 euro in Denmark and Iceland, and 21,000 euro on Norway, but above 30,000 in Germany, Austria and Switzerland.

3. Methods

We use Data Envelopment Analysis (DEA) which is a well-established benchmarking methodology[1]. We analyse different types of models. First, we use DEA to estimate efficiency scores in purely quantitative models with the number of students as output and total expenditure as input. We then investigate whether the estimated efficiency scores are correlated with a large set of indicators of input and output quality. Finally, we estimate models with expenditure per student as input and quality indicators as output (graduation rates, completion rates and expected earnings).

Since DEA and other benchmarking models aim to estimate what is best practice, they are relatively sensitive to single observations. It is therefore important to use good outlier identification techniques, and following the principle of caution it is important to apply aggressive outlier elimination. In all the analyses in this paper, therefore, it has been decided to eliminate production frontier outliers.[2] Thus, countries which have an extraordinary impact on the frontier and thereby on the efficiency estimates of other countries are excluded from the calculations of the efficiency of those other countries.

---

[1]   For discussions of DEA and other frontier analysis methods, see, for example, Bogetoft and Otto (2011) and Bogetoft (2012).

[2]   In doing this we apply two principles based on Banker (1996) and Banker and Chang (2006), which are described in detail in Bogetoft and Otto (2011: 309) and in Appendix A of Bogetoft et al. (2014).

The most obvious challenges that are faced in making international comparisons of the efficiency of the educational sector are: The measurement of resources used, the measurement of services produced and the interactions among the various levels of the educational system.

Resource use can be measured in either monetary or physical units. Physical units include full-time equivalent teacher work years, school buildings, and so on, and calculations in physical units have the advantage of removing the effects of differences in levels of prices, etc. Monetary units, on the other hand, have the advantage of allowing differences in the quality of teachers, buildings, etc. to be taken into account. In this paper we focus on total expenditure (per student) as inputs which have the further advantage that our efficiency measure includes allocative efficiency, i.e. the ability to select an input mix that minimises expenditure.

Production of services is in this paper measured both in purely quantitative terms by the number of students enrolled in educational programmes, and in terms of indicators of the quality of educational production: graduation and completion rates, and expected earnings after completion of education.

The challenge presented by *interactions among educational levels* relates to what students learn at various points in time in the course of their education. It is to be expected that there will be synergy effects, and that *ceteris paribus* there will be greater possibilities for achievement at higher levels in the system if the lower levels have helped students to increase their skills substantially. If, for example, spending on primary/lower secondary education in a country is relatively high, it can be expected that students will complete upper secondary education more easily. Also, the relationships of primary, secondary and tertiary education levels to one another vary somewhat from country to country. For these reasons, one can argue that it may be relevant to include, e.g., expenditure at the primary and lower secondary levels, in addition to expenditure at the upper secondary level, when analysing efficiency of upper secondary education. On the other hand, inputs at the upper secondary level may be assumed to be more important than are inputs at lower levels for outputs at the upper secondary level. Consequently, we analyse models restricted to the upper secondary education level, but we also conduct analyses that include the other levels. Moreover, we also use models in which the contribution made by primary and lower secondary education is measured in approximate form through PISA scores, since these may be interpreted as indicators of the student input quality which the upper secondary level receives from lower levels.

The challenges described above are greatest when the countries concerned are less comparable in general. We can therefore reduce the sources of error significantly if we limit

comparisons to include, for example, only the Northern European countries, where the educational systems, levels of wages and other socioeconomic factors are relatively comparable. The disadvantage of such an approach, however, is that every reduction in the size of the comparison group reduces the potential areas for improvement that can be identified, simply because there are fewer countries in which to discover best practices. This paper presents the results for two comparison groups, designated *All* and *Northern Europe*. The *All* group consists of Western European countries plus certain countries in the rest of the world which are reminiscent of the Nordic countries in various ways, namely New Zealand, Australia, Canada and the USA. The *Northern Europe* group consists of the Nordic countries, Germany, the Benelux countries, the United Kingdom and Ireland (see Table 2). Note, however, that Iceland has not been included in the comparisons, since that country deviates to a very high degree from the other countries in the database used.

Table 2. Primary comparison groups

All		Northern Europe
Denmark	Austria	Denmark
Finland	Switzerland	Finland
Norway	France	Norway
Sweden	Italy	Sweden
Germany	Greece	Germany
The Netherlands	Spain	The Netherlands
Belgium	Portugal	Belgium
Luxembourg	Australia	Luxembourg
United Kingdom	New Zealand	United Kingdom
Ireland	Canada	Ireland
	USA	

4. Quantitative efficiency results: Expenditure and student enrolment

In this section we analyse purely quantitative models in which total expenditure is compared with the numbers of students enrolled at the various levels of education.[3] We present results for five models which are defined in Table 3. The first model covers the entire educational sector, while Model 2 excludes tertiary education. Model 3 additionally excludes the primary level and there focuses on lower and upper secondary levels, while Model 4 is more narrowly focused on upper secondary education. Model 5 is a supplement to Model 3 and examines the differences in the composition of the resources used in the various countries.

Table 3. Quantitative models, i.e. models where output is the number of students, and inputs are PPP-corrected expenditure in 2010

Model	Input: Expenditure	Output: Number of students
1	Total expenditure on primary, secondary and tertiary education	Primary education Lower secondary education Upper secondary education Tertiary education
2	Total expenditure on primary and secondary education	Primary education Lower secondary education Upper secondary education
3	Total expenditure on lower and upper secondary education	Lower secondary education Upper secondary education
4	Total expenditure on upper secondary education	Upper secondary education
5	Expenditure on lower and upper secondary education divided into wages, other operating expenditure and capital expenditure	Lower and upper secondary education

In order to make inputs (expenditure) comparable across countries, expenditure is measured in equivalent US dollars using GDP purchasing power parity (PPP) with the intension to remove most unit cost differences between countries (i.e. a given bundle of education services costs the same across countries). It would be better to use a PPP index based on education services only, but such an index does not exist. Using GDP based PPP is preferable to using standard exchange rates since these primarily reflect prices of internationally traded goods, whereas educational production is predominantly a local service activity.

For all five models we conducted Banker (1996) returns-to-scale tests; see Simar and Wilson (2002) for a discussion of these and other returns-to-scale tests. In the majority of cases the tests

---

[3] In international comparisons of education efficiency it is not common to use DEA models with the number of students as output, although this approach is often used in comparisons of institutions within countries; see e.g. Kirjavainen and Loikkanen (1998).

reject a hypothesis of constant returns to scale. Since we have rather few observations (especially when using the small comparison group of countries) the power of the test is low. Therefore, and in order to be consistent and cautious (i.e., not to exaggerate savings potentials), we report results for all five models assuming non-decreasing returns to scale. Conceptually, the assumption of constant returns to scale could be argued to be reasonable if all countries are large enough that any disadvantages of small-scale systems could safely be ignored; however, we keep to the less restrictive assumption of non-decreasing returns to scale.

The relative efficiencies estimated in the five quantitative models for the two groups *All* and *Northern Europe* (NE) are shown in Table 4. The first row shows the unweighted averages of the efficiencies of the countries concerned.

Table 4. Relative efficiencies as estimated by the quantitative models

Quantitative model	1		2		3		4		5	
Countries	All	NE	All	NE	All	NE	All	NE	All	NE
Average	0.81	0.94	0.82	0.99	0.83	0.95	0.72	0.82	0.95	0.99
Denmark	0.63	0.90	0.66	1	0.88	0.97	0.73	0.91	1	1
Finland	0.97	1	1	1	1	1	0.87	1	1	1
Norway	0.64	0.80	0.64	0.89	0.79	0.79	0.56	0.70	0.94	0.94
Sweden	0.80	0.96	0.81	1	0.88	1	0.76	0.92	1	1
Germany	1	1	1	1[*]	0.89	1	0.56	0.69	1	1
The Netherlands	0.62	0.87	0.65	1	0.56	0.77	0.54	0.66	0.77	1
Belgium			1	1	1	1	0.49	0.61	1	1[*]
Luxembourg							1[*]	1[*]		
United Kingdom	0.85	1	0.80	1	0.86	1	0.74	0.92	1	1
Ireland	0.90	1	0.91	1	1	1	0.95	1	1	1
Austria	0.67		0.68		0.63		0.59		1	
Switzerland	0.63		0.64		0.74		0.71		1	
France	0.74		0.75		0.69		0.54		0.83	
Italy	1		0.94		0.95		0.77		1	
Spain	0.90		0.86		0.92		0.68		1	
Portugal	1		1		1		1		1[*]	
Australia							1			
New Zealand	1		1		1[*]		1			
Canada			0.90		1[*]		0.69		1	
USA	0.62		0.60		0.54		0.50		0.67	

[*] Outlier.

Let us first consider Model 1, which covers the expenditure of the entire education system. When e.g. Denmark is compared with the large group of countries, *All*, its relative efficiency is 0.63. This

is the proportion of Denmark's current expenditure that would be needed if Denmark were to adopt the most efficient best practice from the group of comparison countries. In other words, there exists a combination of countries in the comparison group which could educate the same numbers of students as Denmark at all levels, and which could do this using expenditure which is lower by 1-0.63=0.37, or 37%. If Denmark is compared only with Northern Europe, the savings potential is only 10%.[4] We also see that the potential savings for Denmark are greater than the average potential savings for the other countries. The Danish potential for savings appears to be considerable, since this is a model with four outputs, which from a technical viewpoint is a large number in relation to the number of countries implying that all countries will tend to have relatively high efficiency scores. Relative efficiencies normally decrease with the number of countries in the comparison group, and increase with the number of inputs and outputs. We see here, however, that the level for Denmark is matched only by the levels for Norway, the Netherlands, Switzerland and, to a certain degree, Austria. Finland and Sweden are much more efficient than Denmark and Norway: Finland is fully efficient in the NE comparison group, and nearly so in the All comparison group, whereas Sweden is nearly fully efficient in the NE comparison and has an efficiency score of 0.80 in the All comparison.

When tertiary education is removed from the input and output sides, as is the case in Model 2, the efficiency scores of the Nordic countries are almost the same as for Model 1 in the All comparison, whereas the efficiency scores of Denmark and Norway improve by about 0.1 in the NE comparison (and Denmark becomes fully efficient here). If we also take out primary education, as in Model 3, and focus only on the efficiency of lower and upper secondary education, the efficiency scores of Norway, Sweden and (especially) Denmark increase in the All comparison, whereas Norway's score is reduced by 0.1 in the NE comparison where Finland and Sweden are fully efficient and Denmark nearly so. This reflects the fact, *inter alia*, that expenditure on primary education are relatively high Denmark and Norway and to some extent also in Sweden (see Table 1).

If lower secondary education is removed from the input and output sides, as in Model 4, the efficiency scores of all four Nordic countries decline (particularly for Norway) in the All comparison, and the same is true to a lesser extent for Denmark, Norway and Sweden in the NE

---

[4]   It is generally the case that a larger comparison group tends to lead to lower relative efficiencies. It is however not a logical necessity since we eliminate outliers, and a country may be an outlier in the large group while it is not one in the small group. The small group may therefore contain countries that are excluded from the large group. The outliers are indicated by a star in all the models in the table.

comparison. One possible explanation for this is that New Zealand, which has low expenses at all levels in the educational system (see Table 1), contributes to determining the production frontier in Model 4, whereas in Model 3 it is classified as an outlier (and thus does not contribute to determining the production frontier). The data used in Model 4 and the production frontier are presented in Figure A.1 in the Appendix.

Thus, comparing with the large group of countries, Models 1-4 indicate considerable expenditure-saving potential for Denmark (12-37%), Norway (21-44%) and Sweden (12-24%), whereas Finland is (nearly) fully efficient except in model 4 (for upper secondary education only) where the saving potential is 13%. In the smaller NE comparison group, the expenditure-saving potential is smaller for all four countries: 0-10% for Denmark, 11-30% for Norway, 0-8% for Sweden, and Finland is fully efficient in all models.

When we include information concerning the division of expenditure into wages, other operating expenditure and capital expenditure, as in Model 5, we find that relative efficiencies increase in comparison with Model 3. This is in part because there are now fewer countries in the comparison groups, because of the lack of data for some countries. It is also (and particularly) because there are now three inputs involved instead of just one. The effect of working with three inputs is that the focus is more on technical efficiency rather than efficiency in terms of expenditure. The difference lies in the *allocative efficiency*, i.e. the ability to select an input mix that minimises expenditure. Model 5 does not evaluate the ability of counties to keep total expenditure down by using an appropriate input mix. We observe that according to Model 3, Denmark could save 12% of its total expenditure, while Model 5, in which the existing mix of wages, other operating expenditure and capital expenditure is maintained, suggests that Denmark can make no savings. There is thus evidence to suggest that one possible problem in Denmark could be with the shares of wages, other operating expenditure and capital expenditure in the total expenditure of education. However, it is not possible to state with certainty whether, for example, the proportion of expenditure allocated to wages is too large or too small, since there are countries with optimal allocative efficiency with both larger and smaller proportions of their expenditure spent on salaries than is the case for Denmark. In model 5, Norway's efficiency score is 0.94 (compared to 0.79 in Model 3) and Finland and Sweden are fully efficient for both comparison groups.

5. Explanations for efficiency differences in quantitative models: qualitative differences

Some of the differences in efficiency across countries described above may of course be due differences in the quality of education production, e.g. Model 3 focussing on the secondary level does not take account of the quality of the students entering secondary education, or of those students leaving it. In this section we introduce quality as an explanatory factor, in two ways. First, we examine whether the differences in efficiency between the various countries described in the previous section correlate with important indicators of quality on both input and output sides. Next, we bring in quality indicators such as graduation/completion rates and levels of earnings among students who complete the courses as outputs in DEA analyses, where the input is expenditure per student.

5.1. Output quality indicators

The quality indicators for educational output which we focus on are:

1. Graduation rate: The proportion of a *birth cohort* who complete a given educational programme.
2. Completion rate: The share of *the students who start* a programme who complete it.
3. Expected earnings after completion of an educational programme. This indicator is calculated for upper secondary students as a weighted sum of earnings for three groups: Those who do not complete a programme of upper secondary education; those who complete upper secondary education, but who do not complete tertiary education; and those who complete both upper secondary education and tertiary education. The weights are the proportions of these three groups represented among upper secondary students.
4. Expected earnings after completing the educational programme, corrected for expected rate of employment. This indicator is calculated in the same way as above, but the earnings for the three groups are multiplied by the relevant rate of employment for each group. In this way, the indicator takes into account the facts that the level of education affects both future earnings and future employment. The indicator thus takes into account any trade-off that may exist between high earnings and a high rate of employment.

The earnings used in the calculation of the last two indicators are PPP-corrected, as are inputs, i.e. expenditure (see above). The data and the methods of calculating the indicators are described in detail in Appendix B. It must be emphasised that there is significant uncertainty associated with the calculation of these indicators of quality, and that the methods of calculation and the quality of the data vary from country to country. Similarly, educational systems are to some extent structured differently in different countries. In consequence, the results of these analyses must be interpreted

with caution. It is also important to note that high graduation and completion rates are not necessarily indicators of high quality of the education system – they may reflect low skill requirements for passing exit exams. The third and especially the fourth indicator are probably more unequivocal proxies for the quality of the educational system, and they incorporate the first two indicators, since expected earnings (corrected for employment) depend on the share of students at the upper secondary level who complete this level and the share who eventually complete a tertiary education; see Appendix B for details.

5.2. Correlations between quantitative efficiency and indicators of quality

In this subsection, we examine whether the efficiency differences across countries which we found in the quantitative models of Section 4 correlate with important indicators of quality on both input and output sides. We do this by regressing the efficiency scores from the quantitative models on 25 different quality indicators (one at a time) using the Tobit model, which takes account of the fact that the dependent variables (the efficiencies) are truncated (have values between 0 and 1).[5] We conduct analyses for 25 quality indicators. In addition to the four types of quality indicators described above, we use variables from the OECD database which may serve as indicators of quality on the input side (length of teacher training for various educational levels, and (for the upper secondary and tertiary levels) students' PISA scores). We also examined whether efficiency was correlated with per capita GDP, since one hypothesis might be that countries with high levels of productivity (high GDP per capita) also have high wage costs in the educational sector (producing low efficiency), because the educational sector has to compete for personnel with the highly productive private sector (Baumol's cost disease).

There are only very few significant correlations. This is not entirely surprising given the limitations of the data, including the lack of data for quality indicators for many countries. If we first consider the variables indicating input quality, we would expect that higher quality would be positively correlated with efficiency; but neither the length of teacher training nor PISA scores (which indicate students' scholastic level at the end of lower secondary school) produced significant correlations with efficiency.

On the output side, we might expect that higher quality would require increased expenditure. That would mean, for example, that a higher completion rate – i.e. a lower dropout rate – could

---

[5] The Tobit model is widely used for this type of analysis in the literature, but other methods are used as well, including the one suggested in Papke and Wooldridge (1996); see the discussion in Banker and Natarajan (2008) and Simar and Wilson (2008).

make the teaching more expensive, leading to a lower level of efficiency. We use eight different specifications for completion rates in upper secondary education, and estimate correlations for each of them with the efficiencies calculated in each of the five models of Section 4 (for the comparison group All). The correlation is only significant in two of these 40 estimations, and the sign is different in the two cases. Furthermore, one of the correlations becomes insignificant if we control for GDP per capita in the regression. There are no significant correlations between efficiency and the completion rates for tertiary education, nor between efficiency and graduation rates for upper secondary education. If instead we measure the quality of education at upper secondary level by the share of a birth cohort who later go on to complete tertiary education, there is an indication that higher quality and greater input efficiency may tend to go hand in hand (there is a significant positive correlation in two of the 15 estimations, in both cases for the efficiency estimates in Model 4). Thus, there are no grounds to suppose that the models in Table 4 will show too great an overall potential for savings because of the failure to take into account students' level of preparedness for tertiary education. If quality is measured by level of earnings after completion of a programme, a negative effect is found in Model 1. In other words, the countries where education is relatively expensive are also the ones where the expected earnings on completion of education are high. The findings related to this indicator could suggest that part of the additional expenditure in some countries goes to increasing the quality of the education. However, this correlation is not found when we control for GDP per capita, nor when earnings are corrected for rate of employment (regardless of whether we control for GDP per capita or not).

As expected, GDP per capita seems to be negatively correlated with efficiency (this is the case in three of the five models), indicating that education is generally relatively expensive in countries where productivity is high.

On the whole, these post-analyses of quantitative efficiencies seem to show that the findings that there are overall potential savings to be made are relatively robust. There is little to suggest that efficiency differences between countries are in general explained by quality differences. In this context it should be noted that we have tested 25 variables in each of the five models, and for this reason alone we would expect to find a certain number of false significances.[6] However, including quality indicators directly in the DEA models may well have large effects on the estimated efficiency scores of particular countries compared to the purely quantitative models of Table 4.

---

[6]   Since we tested for significance at the 5% level, we can predict statistically that we will find significance in approximately 6 (5%*25*5) cases, even if there is in fact no true correlation. We found significance in five instances.

The next subsections explore DEA models which include quality indicators.

5.3. The relationship between expenditure per student and indicators of quality

In this section we present the relationships between the four indicators of quality discussed above and expenditure per student in upper secondary education and the educational system as a whole. We hypothesise that, for example, the completion rate for upper secondary education programmes might be particularly strongly affected by expenditure per student enrolled at the upper secondary level, though at the same time the quality of lower secondary education might also have an impact. The quality of lower secondary education can be taken into account by including in the model either expenditure per student at that level or an indicator of students' scholastic level at the end of lower secondary education (which we do in some analyses by including PISA scores). We test a number of combinations of inputs and outputs in the analyses.

Figure 1 displays the relation between the graduation rate for upper secondary education (i.e. the proportion of a birth cohort who complete an upper secondary programme) and average expenditure per student enrolled in primary and secondary education for those countries where the relevant data are available. There are very high graduation rates (over 90%) in Ireland, Finland and the UK, while Denmark and Norway – with 86 and 87% – are on a par with Germany. Sweden's rate is 75%. Luxembourg has by far the highest expenditure per student and the lowest graduation rate, but this is largely attributable to the size of the country. Norway and Denmark have the second and fourth highest level of expenditure per student, while Finland has the lowest. The figure also presents the results from a DEA model based on these data by showing the production frontier (determined by Finland). This will be explained in more detail in the next section, as will the production frontiers shown in Figures 2 and 3, which are discussed below.

Figure 1. Relationship between graduation rates for upper secondary education and expenditure per student in primary and secondary education in 2010
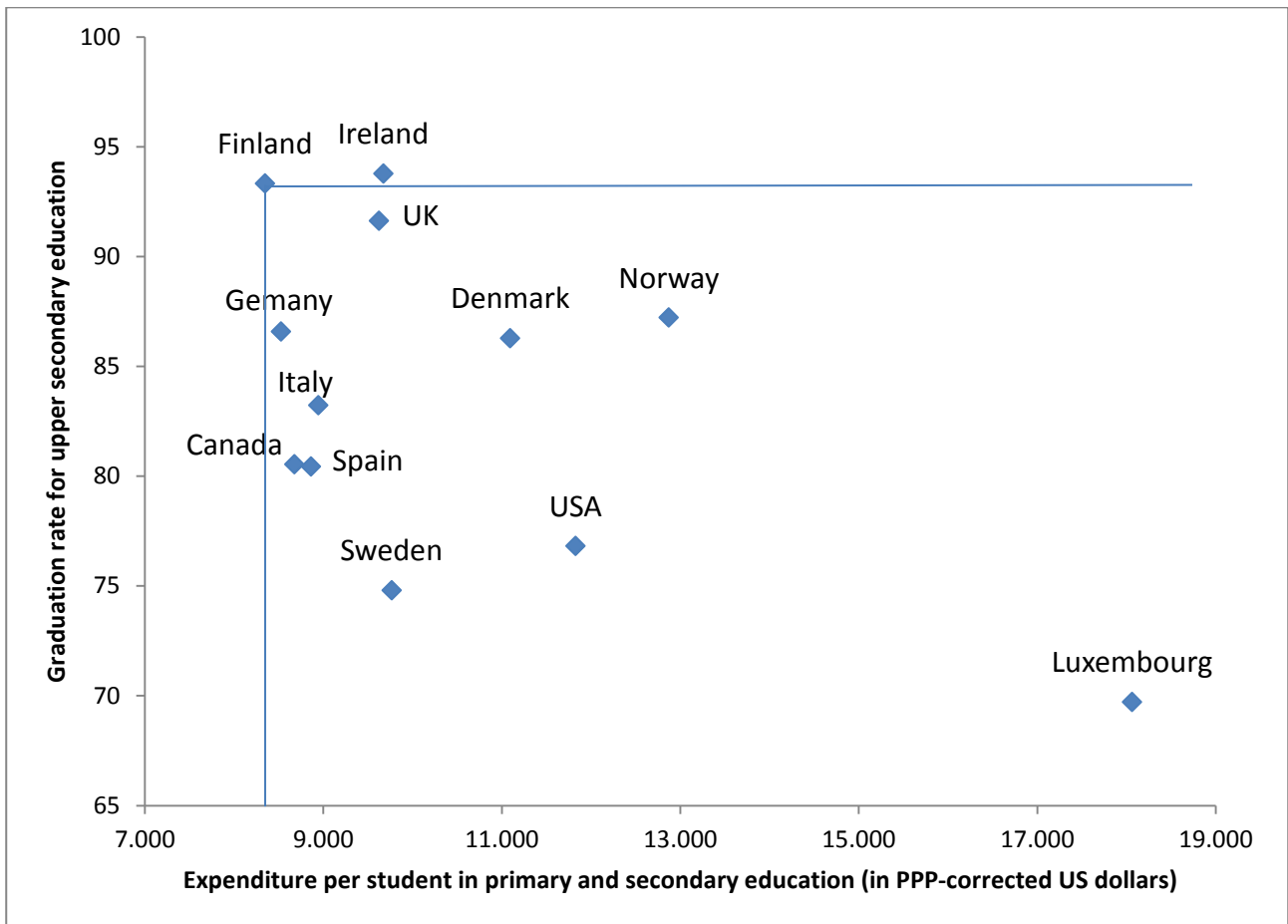

Figure 2 shows the relation between completion rates (within the normal time) for upper secondary education (as a percentage of those who started) and expenditure per student enrolled at this level. Once again, Luxembourg has the highest level of expenditure, Norway the second highest and Finland the lowest; Denmark and Sweden are in the middle of the distribution. Ireland and the USA have the highest completion rates (at around 85%). In Denmark and Norway, the completion rate is relatively low (about 60%), while it is about 70% for Finland and Sweden.

Figure 2. Relationship between completion rates and expenditure per student in upper secondary education in 2010, where the completion rate is the share of the students who start a programme who complete the programme within the normal time
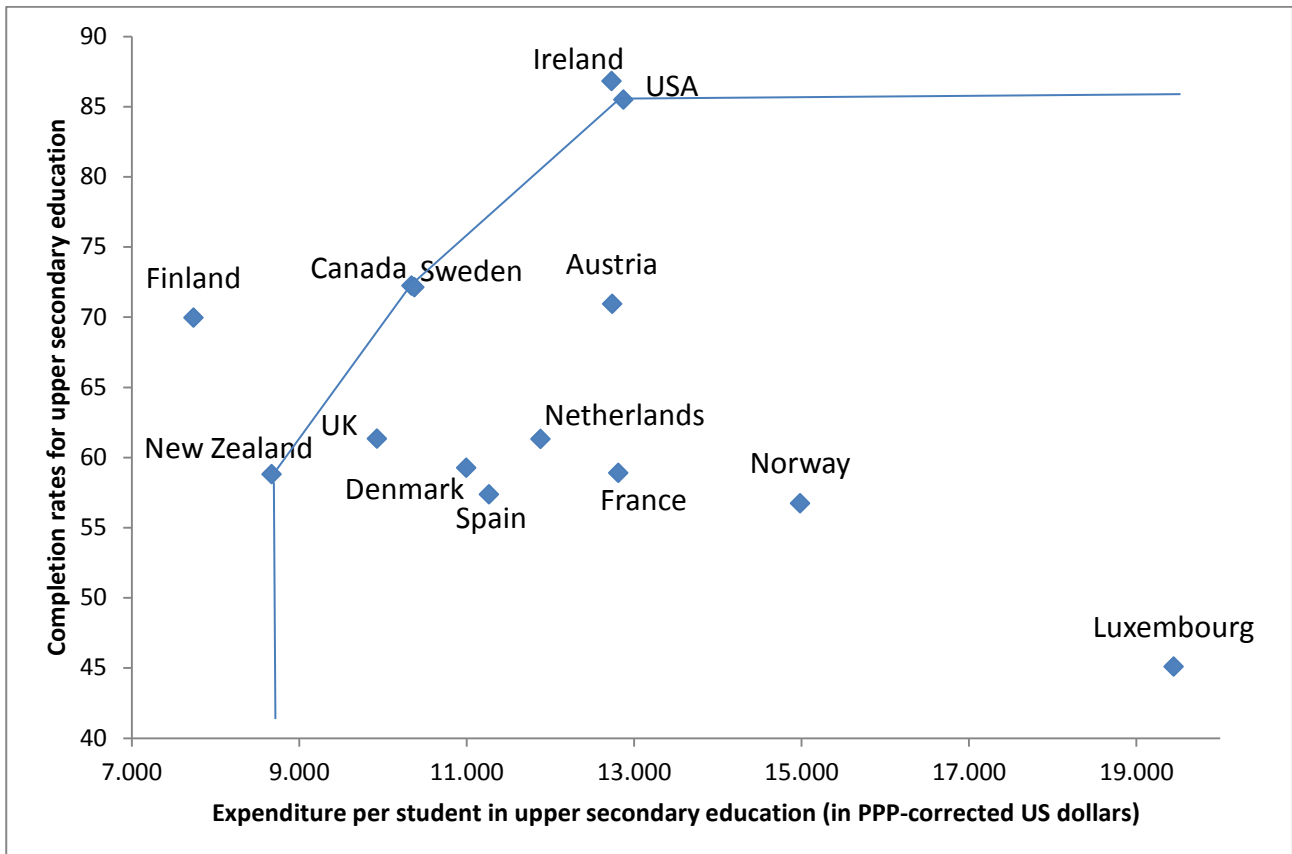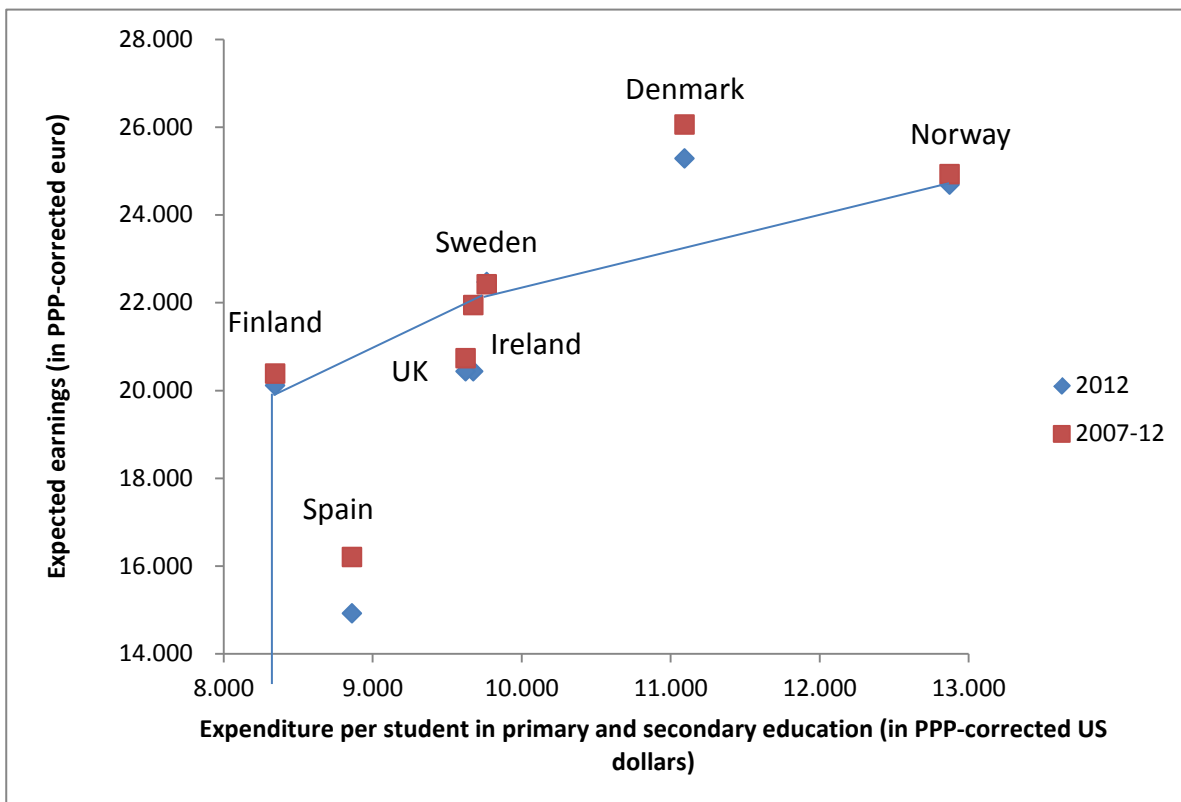
Figure 3 shows the relation between expected future earnings corrected for employment for students in upper secondary education and average expenditure per student in primary and secondary education combined. Denmark ranks first with respect to output, followed by Norway and Sweden. The expected earnings are lower for Denmark, Ireland and Spain when the 2012 rate of employment is used for the correction rather than the average rate of employment for 2007-12, but there are no great differences for Finland, Sweden, Norway and the UK when this alternative correction is applied. This is because the recession affected employment more strongly in Denmark, Ireland and Spain than in the other countries shown here.

Figure 3. Relationship between expected future earnings for students in upper secondary education and average expenditure per student in primary and secondary education in 2010. Expected future earnings are calculated for various educational levels as annual earnings in 2010 multiplied by rate of employment in both 2007-12 (average) and in 2012

5.4. DEA analyses of the relationship between quality and expenditure of resources

This section presents the results from DEA models with the quality indicators described in Section 5.1 as outputs. Data on quality indicators are available from only rather few countries. Therefore,

we only present results for the large group of countries ("All"), and it is necessary to limit the number of other variables in the DEA analyses. We focus on expenditure per student (defined in relation to various elements of the education system) as our central input. This means that we assume constant returns to scale in the relationship between total expenditure and numbers of students. This is a reasonable assumption when comparing countries, since the numbers of students are large in all countries. In addition, we have aggregated parts of the educational system, in that, for example, expenditure per student in primary and secondary school are represented by a single aggregate variable (total expenditure of primary and secondary education divided by the total number of students) instead of e.g. two variables (expenditure per student in primary/lower secondary education and expenditure per student in upper secondary education). In the analyses where output is graduation or completion rates, we consider models with various different specifications of expenditure per student (for the entire education system, for primary and secondary education combined, for secondary education, and for upper secondary education alone).

While an assumption of non-decreasing returns to scale was made in the quantitative models of Section 4 (after rejecting constant returns to scale), the models in this section were estimated using a less restrictive assumption (variable returns to scale) that allows for both increasing and decreasing returns to scale, which is appropriate given that the input is expenditure per student and the output comprises qualitative indicators.[7] As previously, we focus on input-oriented efficiency, in that a country's efficiency score reflects the proportion of its current input that would be necessary to produce its current output if it were as efficient as a combination of the most efficient comparison countries. When we permit variable returns to scale, the country with the highest level of output will be identified as being an outlier, since that country would be able to use any amount of resources (expenditure per student) as input and its score would still not be dominated by that of any of the other countries. This means that the relative efficiencies estimated for other countries are not affected by the country with the greatest output.

---

[7] There are some discussions in the DEA literature about the use of qualitative outputs, or more precisely, ratio outputs. A concern is that such ratios cannot be rescaled as assumed in the non-decreasing or the constant return to scale models since they are restricted to be less than 1. This is one of the reasons we assume variable returns to scale. Note also that we are focusing on input efficiency, i.e. we are looking at potential reductions in expenditure per student, and therefore the output variables serves mainly to restrict the comparison basis. All in all, we therefore argue that the use of ratios in the output specification is applicable.

5.4.1. Graduation rates

Table 5 presents the DEA estimates of countries' relative efficiencies for four models with population graduation rates in upper secondary education as output. Expenditure per student is measured for primary and secondary education in Model 1, for secondary education in Model 2 and for upper secondary education in Model 3. Model 4 is similar to Model 3 except that average PISA scores are included as an extra input. Ireland is an outlier in all the models, because Ireland has the highest graduation rates for upper-secondary education (see Figure 1). Finland is an outlier in Models 3 and 4, due to a combination of low expenditure per student and high values for output, which means that the relative efficiency estimates for other countries are highly dependent on whether or not Finland is included in the calculation.

In Model 1 the production frontier is determined by Finland, since Finland has the lowest expenditure and the highest output (apart from Ireland, which is an outlier). The results for this model are illustrated in Figure 2 above. The analysis indicates that Norway and Denmark could reduce expenditure by (at least) 35 and 25%, respectively, without reducing the graduation rate, if the countries were to become as efficient as Finland. The savings potential for Sweden is 15%. For Norway and Sweden, the relative potential for savings is not affected much when the input is narrowed to expenditure per student in secondary education only or in upper secondary education only; see Models 2 and 3. However for Denmark potential savings are less (19 and 15%) in Models 2 and 3. When PISA results are included as input (Model 4), almost all countries are totally efficient or nearly so (again reflecting the problem with having many variables and few countries in the model).

5.4.2. Completion rates

The population graduation rates are now replaced by completion rates in relation to the number of students who begin an upper secondary education programme, but otherwise the four models in Table 6 are similar to those in Table 5. In Models 1, 2 and 3 of Table 6, where the only input is expenditure per student at different levels of the educational system, Finland is fully efficient (or nearly so), Sweden has efficiency scores between 0.89 and 1, whereas the efficiencies of Norway and Denmark are very low (0.63-0.65 and 0.75-0.85, respectively). The results for Model 3 are presented in Figure 2 above. If PISA results are included as inputs (Model 4) then again all countries have relative efficiencies that are close to 1.

The completion rates used in Table 6 and Figure 2 are defined in terms of completion within the normal duration of the programme. However, robustness checks show that when completion of an upper secondary programme is counted as long as it is within the normal duration of the programme plus two years, relative efficiencies for the Nordic countries do not change much.

Table 5. Relative efficiencies from models where the output is graduation rate measured as the share of a birth cohort that completes an education.

Model	1	2	3	4
Expenditure per student	Primary and secondary education	Secondary education	Upper secondary education	Upper secondary education
Additional input				PISA scores in reading and math
Average	0.84	0.85	0.84	0.99
Denmark	0.75	0.81	0.85	0.99
Finland	1	1	1*	1*
Norway	0.65	0.64	0.63	0.98
Sweden	0.85	0.89	0.87	0.98
Germany	0.98	0.96	0.83	0.98
Luxembourg	0.46	0.46	0.47	1
UK	0.87	0.89	1	1
Ireland	1*	1*	1*	1*
Italy	0.93	0.98	1	1
Spain	0.94	0.88	0.81	1
Canada	0.96	0.99	0.88	0.93
USA	0.71	0.71	0.71	0.99

*Outlier

Table 6. Relative efficiencies from models where the output is student completion rates measured as the share of the students who begin an education programme who go on to complete it within the normal duration of the programme

Model	1	2	3	4
Expenditure per student	Primary and secondary education	Secondary education	Upper secondary education	Upper secondary education
Additional input				PISA scores in reading and math
Average	0.83	0.81	0.84	0.99
Denmark	0.68	0.72	0.79	0.98
Finland	1	0.99	1*	1*
Norway	0.58	0.57	0.58	0.97
Sweden	0.89	0.89	1	1*
Netherlands	0.77	0.69	0.76	0.95
Belgium	0.84	0.81		
Luxembourg	0.41	0.41	0.45	1
UK	0.80	0.81	0.90	1
Ireland	1*	1*	1*	1
Austria	0.72	0.71	0.80	1
France	0.84	0.74	0.68	0.97
Spain	0.85	0.79	0.77	1
New Zealand	1	1	1	1
Canada	1	1	1	1
USA	1	1	1	1

*Outlier

While we would ideally like to include many inputs and outputs simultaneously in the analyses, there are relatively few countries that it is relevant to compare with the Nordic countries and for which the necessary data are available. In a DEA analysis, where the underlying assumptions about the production process are not restrictive, virtually all countries will be categorised as efficient if the number of inputs and outputs is high relative to the number of countries. This being the case, we cannot tell whether the fact that especially Denmark and Norway appear to be very inefficient in many of the simple models with few inputs and outputs is due to real inefficiency, or to the fact that the models are very simple, with many factors that are not taken into account.

5.4.3. Expected earnings

Table 7 shows the results found when output is expected earnings after completion of an educational programme for students in upper secondary education, and input is average expenditure per student in primary and secondary education. Even though there is only one input and one output, very few countries can be included, on account of a lack of data. The output in Model 1 of Table 7 is expected future earnings without taking into account expected future rates of employment. These expected earnings are, as described in Section 5.1 and Appendix B, a weighted average of annual earnings in 2010 for various levels of education. The output in Model 2 is expected earnings corrected for expected rate of employment, i.e. expected earnings are the weighted averages of annual earnings in 2010 for various levels of education multiplied by the rates of employment for these levels of education. Rates of employment vary greatly from year to year, and we have therefore based our calculations on average rates of employment for various levels of education in 2007-2012. Denmark is an outlier in these estimates, because the expected earnings are higher than for other countries, and this also applies when a correction is made for rate of employment. Finland and Norway are also fully efficient in this context, as is Sweden when correction is made for employment rate. Without correction for employment, the efficiency score of Sweden is 0.85. The data underlying the results of Model 2 are shown in Figure 3 above, which also shows the results when average employment rates for 2007-2012 are replaced by employment rates for 2012 alone.

Table 7. Efficiency scores for models where the output is earnings in 2010 or earnings in 2010 multiplied by rate of employment in 2007-12, and where the input is expenditure per student in primary and secondary education in 2010. Both inputs and outputs are PPP corrected.

Model	1	2
Output	Earnings	Earnings × rate of employment
Average	0.95	0.97
Denmark	1*	1*
Finland	1	1
Norway	1	1
Sweden	0.85	1
United Kingdom	0.87	0.89
Ireland	1	0.97
Spain	0.94	0.94

*Outlier

5.5. Comparisons with results of other DEA studies

Not many DEA analyses exist that focus on international comparisons of upper secondary education. Verhoeven et al. (2007) use DEA to analyse the relation between PPP-corrected expenditure per student in secondary education and upper secondary graduation rates, basing their research on data from *Education at a Glance 2006*. Their analysis comprises more countries, including Eastern European countries, and their results indicate that the potential savings for Denmark, Finland and Sweden are very considerable (Verhoeven et al., 2007: Figure 6); Norway has the largest expenditure per student among the Nordic countries, but is on the efficient production frontier due to a recorded graduation rate of about 100%, presumably caused by data issues for that year. If the comparison countries are limited to the same group that we have included, their estimated savings potential for Denmark is approximately the same as ours (20-25%), whereas their estimates for Finland and Sweden are larger (about 16 and 25%).[8] Verhoeven

---

[8] These results are obtained by visual inspection of Figure 6 in Verhoeven et al. (2007) where a restriction of the sample to our comparison group implies that the efficient frontier would be determined by the observations for New Zealand, Ireland and Germany.

et al. (2007) also contains a corresponding analysis for tertiary education, which indicates that Norway and Denmark, and especially Sweden, have substantial savings potentials, also after restricting the comparison group to the one we use (see their Figure 7). Afonso and St. Aubyn (2005) carried out a DEA analysis where the focus was on the last year of lower secondary school. The output used was PISA scores, and the inputs were the number of teachers per student and time devoted to teaching. In this analysis, Finland and Sweden are fully efficient, whereas Denmark ranks eleventh among 17 OECD countries, with a potential for savings of around 14%; Norway is not included in the analysis. In a similar analysis with corrections for GDP per capita and parents' education (Afonso and St. Aubyn 2006), Finland, Sweden, Denmark and Norway rank 8, 15, 23 and 25, respectively, among 25 OECD countries.

6. Robustness check: analyses for year 2000

As discussed above, the results concerning the relative efficiencies of various countries must be interpreted with caution. One reason is that the analyses are based on data for a single year, namely 2010 (though with some supplementary data for 2011 for completion rates, and for 2007-2012 for employment rates). However, enrolment, expenditure and quality indicators can vary over time in any given country. The analyses described above of the relationship between expenditure per student and quality indicators implicitly assume that inputs and outputs remain relatively stable over time for a given country. For example, the completion rate for upper secondary education in a given year is actually linked to the resources expended on the students concerned both in that year and in the previous three or four years, and to the resources expended in primary/lower secondary school even further back in time. Similarly, the earnings and employment rates which students can expect after they have reached their final level of education depend on the quality of the educational system over many previous years. It is very difficult (not to say impossible) to perform satisfactory analyses that take this time element into account, since data would be missing for many countries for some of the relevant years. Instead, we describe in this section the results of a more modest robustness check, in which we conduct analyses for the period around the year 2000 instead of the period around 2010 used in the main analysis. Naturally, we have been obliged to limit these additional analyses to the models for which the key data are available for the year 2000 or around that time (e.g., there are no OECD data for completion rates around year 2000). Below we summarize the main results of these robustness checks (more details can be found in Bogetoft et al., 2014).

For the purely quantitative models corresponding to Models 1-4 of Table 4 the results are rather robust for Finland and Sweden, but not for Denmark and Norway. Finland is still fully efficient in most models; Sweden's efficiency scores range 0.72-0.96, similar to the 2010 results; Norway's efficiency scores were better in 2000 (range 0.63-1.00), whereas Denmark's were considerably worse (range 0.56-0.76).

For the models with upper secondary education graduation rates as output (corresponding to Models 1-3 of Table 5), the results are very different in 2000 compared to 2010. In 2000 Norway is fully efficient in most cases (whereas Norway rank lowest among the Nordic countries in 2010); Finland which were fully efficient in 2010 has much lower scores in 2000 (0.75-0.85); and the scores for Denmark in particular, but also for Sweden, were also much lower in 2000 (0.48-0.53 and 0.64-0.82, respectively). One of the reasons why these results are so unstable is probably the very crude way in which graduation rates are calculated in the OECD data; see Appendix B.

Results for expected earnings corrected for employment rates are very similar for 2000 compared to 2010: All the four Nordic countries are fully efficient, except Sweden for which the efficiency score is 0.96.

7. Conclusion

With the aim of investigating the relative efficiency of education production in the Nordic countries with particular emphasis on the upper secondary level, we have defined a group of comparable OECD countries and used the DEA approach to calculate efficiency scores in a range of models, which differ by the specification of inputs and outputs. In quantitative models, where the input is total expenditure and outputs are enrolment at different levels of the educational system, Finland is fully efficient in most specifications, whereas the savings potentials of Sweden, Denmark and Norway are in the range of 0-26%, 0-37% and 11-44%, respectively (i.e. Sweden and Denmark are fully efficient in a few specifications with a narrow North European comparison group).

In models which have as output graduation or completion rates for upper secondary education and as input expenditure per student (Models 1-3 in Tables 5 and 6), Finland is fully efficient (or nearly so), whereas the estimated potential savings are large for Norway (35-43%) and Denmark (15-32%) and more moderate for Sweden (0-15%), depending on the definitions used of graduation/completion rates and expenditure per student. When the countries' PISA scores are included, countries such as Denmark, Norway and Sweden, which perform relatively poorly on the PISA tests, make up ground on countries such as Finland and New Zealand, which perform well,

and the differences in relative efficiency are then rather small; there is, however, a problem with these analyses, in that the small number of observations relative to the number of variables tends to increase the estimates of relative efficiency.

All four Nordic countries are found to be fully efficient in analyses where the output is expected earnings after completion of educational programs corrected for employment and the input is expenditure per student. Thus, the main conclusions are the following. First, Finland is very efficient in terms of education production in all model specifications. The other Nordic countries, and especially Denmark and Norway, are far from being efficient in purely quantitative models and in models where graduation or completion rates are used as indicators of the quality of the education system; thus, a policy implication from these models might be that there are large savings potentials in these countries. However, graduation and completion rates are very crude indicators of the quality of education production; they say nothing about the qualifications or productivity of the students who obtain various levels of education. Therefore, we introduced a more comprehensive quality indicator, namely expected future earnings of students enrolled in upper secondary education. In models using this indicator as output, all Nordic countries become fully efficient, i.e. they are more efficient than the other countries in the comparison group. However, it is important to note that, because of data issues, there are very few countries in this analysis. Also, even if this analysis indicates that the Nordic countries are efficient relative to other countries, it would require cost-benefit analyses, which are beyond the scope of this paper, to inform whether the benefits of higher future earnings outweigh the higher expenditure on education in the Nordic countries.

International comparisons of efficiency in education production are very relevant, not least in light of the growing internalisation of education, but they are also very challenging. There are significant differences among countries with regard to the structure of their education systems, including whether or not pupils are divided up into different trajectories within an educational level, and there may – and probably will – be significant spill-over effects between levels. Such institutional differences contribute to the difficulties in setting up datasets with variables that are truly comparable across countries. Thus our results should be interpreted cautiously. Furthermore, the small number of relevant comparison countries and missing data problems imply that it is typically only possible to include very few variables at a time in the analysis. We have conducted a range of different estimations using different specifications of inputs and outputs and different comparison groups, and the results do indeed depend on the specification.

Results using data for 2010 (as in our main analysis) are for some model specifications rather different from results obtained for similar models using data for 2000. One reason for this may be that relative efficiencies have changed over time, but it may also be due to data issues, for instance problems in calculating valid measures of graduation rates. However, results are very similar in 2000 and 2010 when expected earnings corrected for employment rates are used as output variable. In future research it would be interesting to find data enabling calculation of this output quality indicator for a larger group of countries and, if possible, to use longitudinal data to calculate earnings and employment rates explicitly for the relevant cohorts.

In future research it would also be interesting to apply the approach of Afonso and St. Aubyn (2006) and correct the estimated efficiency scores for important non-discretionary inputs such as parental education. In this case the corrected efficiency scores will to some extent reflect changes in efficiency over time instead of the level of efficiency at a given point in time. This is perhaps most apparent, when the variable for parental education is similar to the educational output variable used in the DEA model, e.g. if parental education is measured by the share of the parental generation having at least an upper secondary education (as in Afonso and St. Aubyn, 2006), i.e. the upper secondary graduation rate of these cohorts, and the output variable is the upper secondary graduation rate of the younger cohorts.
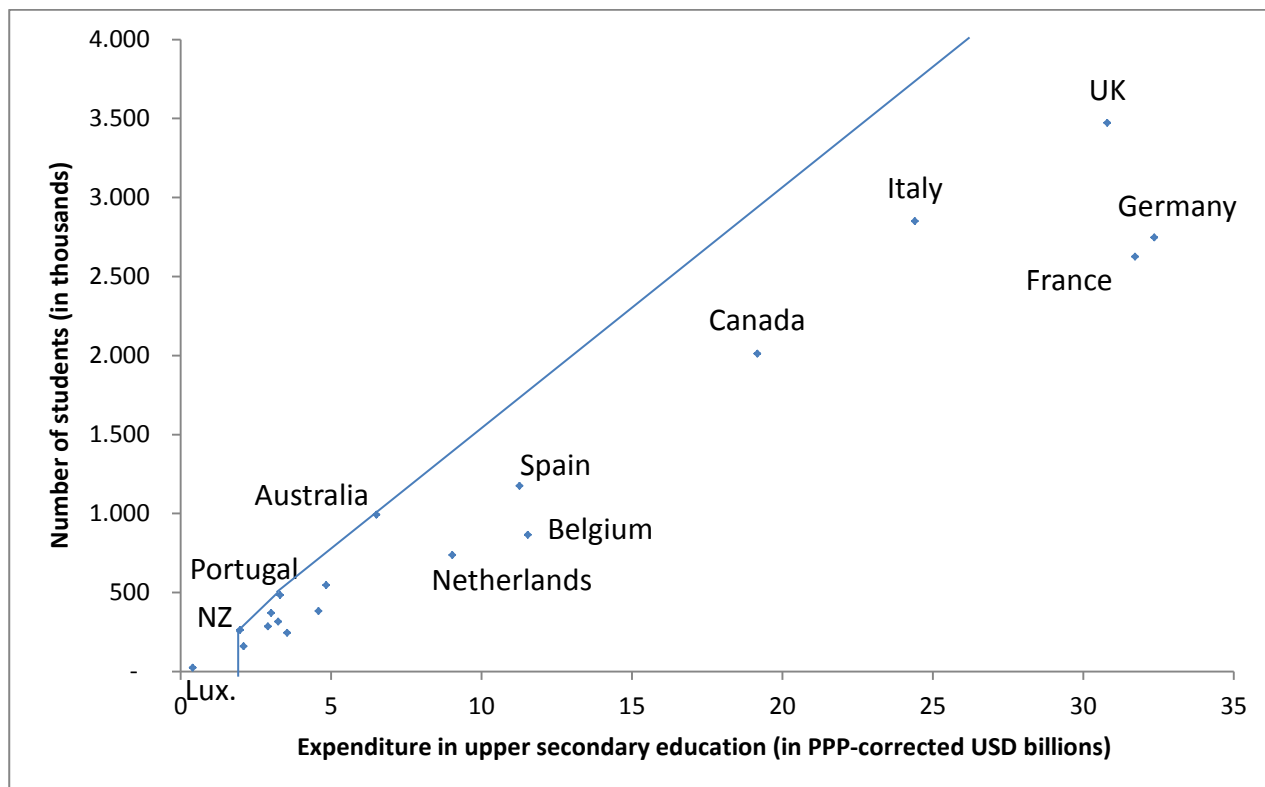
Appendix A. Illustration of efficiency in quantitative models

Figure A.1 illustrates the relationship between the number of students in upper secondary education and the total expenditure of upper secondary education (measured in PPP-corrected US dollars). Each dot in the figure represents the data for one country. Large countries with many students and high total expenditure lie on the right-hand side of the figure (and near the top), while small countries are on the left-hand side. (The USA is not included on the figure, since it is so much larger than other countries that it would be difficult to identify other data points if it was shown.) The line shows the production frontier, i.e. the maximum level of efficiency found, plotted on the assumption that there are non-decreasing returns to scale (i.e. that the unit costs do not increase with the number of students). New Zealand (NZ in the figure), Portugal and Australia lie on the production frontier and are thus the most efficient countries. The assumption of non-decreasing returns to scale means here that the production frontier is determined by unit costs in Australia for all points to the right of and above the level of Australia. The Nordic countries are represented by the dots close to New Zealand, e.g. Denmark is represented by the dot positioned immediately to the right of New Zealand (with approximately the same number of students). Luxembourg is a special case because of its size, and is an outlier. A country's saving potential (and its relative efficiency score) is indicated by its (relative) horizontal distance from the production frontier. The United Kingdom, for example, spends approximately USD 31 billion, and its horizontal distance from the production frontier is around USD 8 billion. Its absolute saving potential is thus USD 8 billion, and relative to its actual expenditure its saving potential is $8/31 = 26\%$; in other words, its efficiency is 74%.

Figure A.1. The relationship between expenditure in upper secondary education and numbers of students in upper secondary education in 2010 (graphical representation of the results from Model 4 in Table 4)

Appendix B. The four indicators of output quality: Data and specifications

1. Graduation rates (completion rates in relation to the *population*)
   o In the case of upper secondary education, this rate is measured as the share of a birth cohort of the relevant age that completes a programme of secondary education. In the OECD's data, this is calculated by dividing the number of students who graduate from upper secondary education in a given year (2010) by the total number of people in the population of the typical age for completion of upper secondary education (OECD 2012, Annex 3, p. 23). The calculation is inevitably affected by the fact that there may be great variation in the actual age of people graduating from upper secondary education and in the size of age cohorts.

- In the case of tertiary education, this rate is measured as the share of a birth cohort of the relevant age made up of students who graduate *for the first time* from a programme of tertiary education, regardless of the age at which they graduate. The calculation is based on the number of students graduating from a programme of tertiary education in a given year (2010) and their age distribution (OECD 2013, p. 55). There is a special problem in calculating the graduation rate for tertiary education with respect to overseas students. If a large proportion of the young people in a given country take their tertiary education in a foreign country, the graduation rate will be an underestimate (e.g. in the case of Luxembourg), whereas it will be an overestimate in the case of a country such as the UK, where the number of overseas students is high (even calculated net of UK students going abroad to study).

2. Completion rates (the share of the students *who started a programme* who go on to complete it)
   - Upper secondary programmes: For over half the countries in the study, the completion rate for upper secondary education was calculated using a true cohort method, i.e. students who began an upper secondary programme in a given year were followed up on an individual basis to discover how many completed the programme within its normal duration, and within the normal duration plus two years. For the remaining countries, the completion rate was calculated as the number of students completing an upper secondary programme in a given year (2010) divided by the number who began the programme N years previously, where N is the normal duration of the programme (OECD 2012, Annex 3, pp 27-8). This method of calculation takes into account changes in the size of the relevant age cohort, but does not take into account whether or not students completed the programme within the normal time.
   - Tertiary education: For around half the countries in the study, the completion rates were calculated using a true cohort method. All the individuals who began a tertiary programme in a given year were traced, and it was determined how many had later completed that or another programme of tertiary education. For the remaining countries, the completion rate was calculated as the number of students completing a tertiary programme in a given year (2011) divided by the number who began the programme N years previously, where N is the normal duration of the programme (OECD 2013, p. 70).
3. The earnings rates used in the main analysis are for 2010 and are based on data from Eurostat (Eurostat 2014: Dataset earn_ses10_30). Earnings for each of the three levels are based on average earnings for people in employment with ISCED educational levels 0-2, 3-4 and 5-6 (where, for example, the distribution of the population across levels 3 and 4 provides the basis for combining the earnings for level 3 and for level 4). More precisely, the expected earnings for students in upper secondary education (EUS) are determined as:

$$EUS = (1 - CRUS) * E02 + CRUS * (1 - GRT)E34 + CRUS * GRT * E56$$

where *CRUS* is the completion rate for upper secondary level; *E*02, *E*34 and *E*56 are, respectively, average earnings for individuals with compulsory schooling, upper secondary education and tertiary education as their highest completed educational level; and *GRT* is the graduation rate for tertiary education related to students in upper secondary level (i.e. graduation rate tertiary / graduation rate upper secondary).

4. The rates of employment used are based on Eurostat (2014: Dataset lfsa_ergaed) and in the main analysis are an average for the years 2007-12. In a supplementary analysis in the study, the rates are based on data for 2012 alone. Expected earnings are corrected for rates of employment calculated in the same way as EUS above, except that the earnings for the various levels of education (*E*02, *E*34 are *E*56) are replaced by earnings multiplied by rate of employment for the corresponding educational level, i.e. *E*02**emp*02, *E*34**emp*34 and *E*56**emp*56, where *emp*02, *emp*34 and *emp*56 are rates of employment.

References

Afonso, A., St. Aubyn, M, 2005. 'Non-parametric approaches to education and health efficiency in OECD countries'. Journal of Applied Economics, 8 (2), 227-246.

Afonso, A., St. Aubyn, M., 2006. 'Cross-country efficiency of secondary education provision – A semi-parametric analysis with non-discretionary inputs'. Economic Modelling, 23 (3), 476-491.

Banker, R.D., 1996. 'Hypothesis tests using Data Envelopment Analysis'. Journal of Productivity Analysis, 7 (2-3), 139-160.

Banker, R.D., Chang, H., 2006. 'The super-efficiency procedure for outlier identification, not for ranking efficient units'. European Journal of Operational Research, 175 (2), 1311-1320.

Banker, R.D., Natarajan R., 2008. Evaluating contextual variables affecting productivity using data envelopment analysis. Operations Research 56 (1), 48-58.

Barros, C.P., Guironnet, J.P., Peypoch, N., 2011. Productivity growth and biased technical change in French higher education. Economic Modelling 28, 641-646.

Bogetoft, P., 2012. *Performance Benchmarking – Measuring and Managing Performance*. New York: Springer.

Bogetoft, P., Heinesen, E., Tranæs, T., 2014. *The efficeincy of educational production: A comparison of Denmark with other OECD countries.* Study Paper No. 71. Copenhagen: Rockwool Foundation Research Unit.

Bogetoft, P., Otto, L. 2011. Benchmarking with DEA, SFA, and R. New York: Springer.

Bogetoft, P., Wittrup, J., 2011. Productivity and education: Benchmarking of elementary and lower secondary schools in Denmark. Nordic Economic Policy Review 2, 257-294.

Coco, G., Lagravinese, R., 2014. Cronyism and education performance. Economic Modelling 38, 443-450.

Eurostat, 2014. Eurostat database, http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/search_database. Data for earnings and employment rates by education. Accessed January-April 2014.

Giménez, V., Prior, D., Thieme, C., 2007. Technical efficiency, managerial efficiency and objective-setting in the educational system: an international comparison. The Journal of the Operational Research Society 58 (8), 996-1007.

Gupta, S., Verhoeven, M., Tiongson, 2002. The effectiveness of government spending on education and health care in developing and transition economies. European Journal of Political Economy 18, 717-737.

Hanushek, E.A., Kimko, D.D., 2000. Schooling, labor-force quality, and growth of nations. American Economic Review 90 (5), 1184-1208.

Hanushek, E.A., Luque, J.A., 2003. Efficiency and equity in schools around the world. Economics of Education Review 22, 481-502.

Herrera, S., Pang, G., 2005. Efficiency of public spending in developing countries: An efficiency frontier approach. World Bank Policy Research Working Paper No. 3645.

Kirjavainen, T., Loikkanen, H.A., 1998. Efficiency differences in Finnish senior secondary schools: An application of DEA and Tobit analysis. Economics of Education Review 17 (4), 377-394.

Lu, W.M., 2012. Intellectual capital and university performance in Taiwan. Economic Modelling 29, 1081-1089.

Minnigou, E.W., Vierstraete, V. 2013. Households' living situation and the efficient provision of primary education in Burkino Faso. Economic Modelling 35, 910-917.

OECD, 2006. Education at a Glance 2006.

OECD, 2012. Education at a Glance 2012.

OECD, 2013. Education at a Glance 2013.

Papke, L.E., Wooldridge, J.M., 1996. Econometric methods for fractional response variables with an application to 401(k) plan participation rates. Journal of Applied Econometrics 11: 619-632.

Simar, L., Wilson, P.W., 2002. Non-parametric tests of returns to scale. European Journal of Operational Research 139, 115-132.

Simar, L., Wilson, P.W., 2007. Estimation and inference in two-stage, semi-parametric models of production processes. Journal of Econometrics 136 (1), 31-64.

Simar, L., Wilson, P.W., 2008. Statistical inference in non-paramettric frontier models: recent developments and perspectives. In Fried, H.O., Knox-Lovell, C.A., Schmidt, S. (Eds.): The Measurement of Productive Efficiency and Productivity Growth. Oxford University Press, Oxford.

Sutherland, D., Price, R.W., Joumard, I., Nicq, C., 2007. Performance indicators for public spending efficiency in primary and secondary education. OECD Economics Department Working Papers, No. 546, OECD Publishing.

Verhoeven, M., Gunnarsson, V., Carcillo, S., 2007. Education and Health in G7 Countries: Achieving Better Outcomes with Less Spending. IMF Working Paper WP/07/263. International Monetary Fund.

Zoghbi, A.C., Rocha, F., Mattos, E., 2013. Education production efficiency: Evidence from Brazilian universities. Economic Modelling 31, 94-103.